DIAGNOSTIC REASONING MODEL
VALIDATION IN DIGESTIVE ENDOSCOPY.

JM Cauvin[1], C Le Guillou[2], B Solaiman[2], M Robaszkiewicz[3], H Gouérou[3], C Roux[4].

[1]LATIM, Medical Information Department. University Hospital. 29609 Brest - France.
[2]LATIM, Image and Information Treatment Department. ENSTB. 29285 Brest - France.
[3]LATIM, Hepato-Gastroenterology Department. University Hospital. 29609 Brest - France
E-mail : jean-michel.cauvin@univ-brest.fr

Abstract: **The development of a computer-assisted diagnostic system in digestive endoscopy implies to understand the reasoning process of endoscopists. The aim of this study is to validate a reasoning model and a knowledge base previously defined. Eight endoscopists have participated to a diagnostic test including 5 video-sequences and using a "thinking aloud" method. The physician discourse was compared with the knowledge base content. Forty requests, built with indices extracted from endoscopist talks were submitted to the diagnostic research system.**

The main results show that the endoscopists evoke all the information types defined in the knowledge base. Endoscopists well distinguish the endoscopic findings and diagnoses. They combine data-driven and hypothesis-driven approaches during the diagnostic process. Faced to complex diagnostic problem solving, they promote an analytical approach. The 40 requests have led to the correct finding class, as unique or first response in 25, second or other rank response in 12, and failed to recognize the lesion in 3 cases.

These results are consistent with the reasoning model and the knowledge base content. They suggest several ways for improving the diagnostic aid system, the research approach as well as the user interface.

Keywords: **Medical reasoning, endoscopy, validation.**

I. INTRODUCTION

Endoscopy has modified the diagnostic approach of digestive tract diseases in a radical way, because it permits direct observation of digestive lesions with low invasive device and low risk for patients. With videoendoscopy, sequences of images can be stored more easily to constitute a medical referring document of the medical record file. The purpose of our project is to build an atlas of indexed endoscopic lesions that could be used in computer-assisted diagnostic, as referring data. Development of "intelligent" tools, which can retrieve referring images similar to an observed lesion, implies to understand the diagnostic reasoning. The aim of this study is to validate with endoscopists, the suggested reasoning model and the nature of the knowledge mobilized in diagnostic endoscopy as they were defined in previous studies [1,2].

In a first part, we will briefly describe the reasoning model and the knowledge base that has been deduced from literature review and endoscopist expertise. In the second part, we will report the test modalities and results, which will be discussed in the last part of the paper.

II. DIAGNOSTIC REASONING MODEL

The decision of digestive endoscopy is always justified by a medical context. The nature of this medical context leads to distinguish two reasoning approaches, which refer to three levels of medical knowledge (fig. 1). In one hand, the medical context is not specific: systematic exploration of the organ permits to focus on lesions, according to abnormal variation of color, relief or anatomical repairs. Medical knowledge of their characteristics leads to diagnostic of these elementary objects. Association of lesions compared with knowledge of gut diseases leads to generation of diagnostic hypothesis [3]. In the other hand, specific context leads to formulate diagnostic hypothesis very earlier, and combination of endoscopic lesions whose association confirms the diagnosis. The endoscopist explores more closely gut regions where these lesions are usually found.
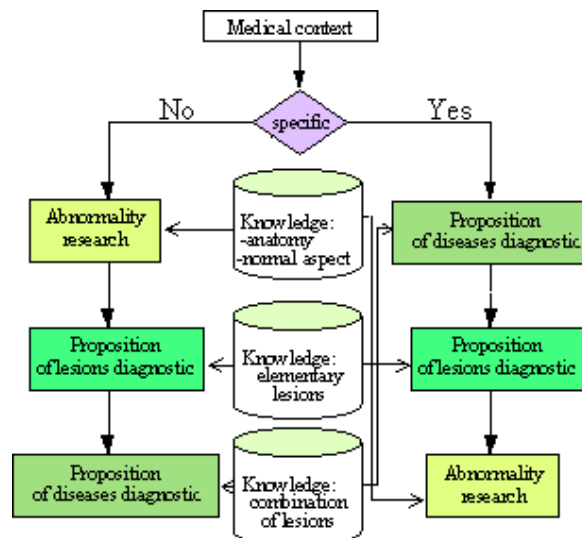

Fig. 1: Diagnostic reasoning model in digestive endoscopy.

During endoscopic examination, the two approaches are not exclusive [4]. In the later, the rule of systematic examination persists and unsuspected lesions can be discovered. In the former, generation of diagnostic hypothesis can lead to look for specific elementary lesions whose absence or presence confirms the endoscopic diagnostic.

Report Documentation Page

Report Date 25 Oct 2001	Report Type N/A	Dates Covered (from... to) -

Title and Subtitle		Contract Number
Diagnostic Reasoning Model Validation in Digestive Endoscopy		
		Grant Number
		Program Element Number

Author(s)		Project Number
		Task Number
		Work Unit Number

Performing Organization Name(s) and Address(es)	Performing Organization Report Number
LATIM Medical Information Department University Hospital 29609 Brest - France	

Sponsoring/Monitoring Agency Name(s) and Address(es)	Sponsor/Monitor's Acronym(s)
US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Oct 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification unclassified	Classification of this page unclassified

Classification of Abstract unclassified	Limitation of Abstract UU

Number of Pages
4

III. Knowledge Representation

The medical reasoning during endoscopy can be compared to the interpretation of the scene with objects where the scenes are the endoscopic diagnoses and the objects are the set of gathered information before and during the examination: demographic data, examination reasons, endoscopic findings and their spatial relationships (fig. 2). Classes of endoscopic findings and diagnoses constitute the two decision spaces of the diagnostic process.

The most usual endoscopic reasons for examination, findings and diagnoses, have been listed in the Minimal Standard Terminology for a computerized endoscopic database, which has been used as a classification of endoscopical scenes and objects [5]. Spatial relationships between objects have been deduced from endoscopist expertise.
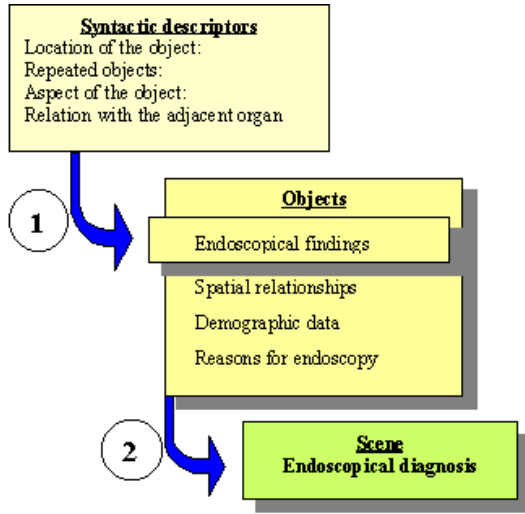

Fig. 2. Problem solving representation in digestive endoscopy.
1: endoscopic finding decision. 2: endoscopic diagnosis decision.

The syntactic descriptors of endoscopical anomalies and the content of each descriptor cannot be retrieved in extension in this terminology. Based on endoscopist expertise, 19 elementary characteristics (primitives) of objects have been identified (Table 1). Most of them can be extracted from the observed image. Others call for different sense stimulus: for example, consistency is deduced from palpation of the lesion with an endoscopical device.

A simple object is described by all these attributes. A complex or "non-homogenous object" is composed of 2 or more simple objects fitted together. Each of them has its attributes, and a spatial inclusion relation links them. The main object has the attributes described in table 1. The 9 sub-object characteristics are selected in the "aspect of the object" descriptors. The content of each descriptor have been identified and 128 distinct values have been defined for the 19 descriptors of the main object and 65 for the sub-object.

From a medical point of view, it is possible to describe the classes of objects and scenes in an intensive way according to defined descriptors. This approach has led to an *a priori* description of 97 endoscopical findings and 158 endoscopical diagnoses.

The retrieval process, based on weighted relationships between syntactic descriptors and findings, and between objects and scenes has been previously described [2].

TABLE 1:
Syntactic descriptors of endoscopical lesions

Location of the object:
    Anatomic location (longitudinal),
    Position in the organ (axial),
    Distance from the teeth.
Repeated objects:
    Number of identical objects,
    Spatial organization.
Aspect of the object:
    Shape and edge,
    Dominant color and color regularity,
    Relief and regularity of relief,
    Sizes: height, width and thickness,
    Motility,
    Effect of insufflation,
    Consistency.
Relation with the adjacent organ:
    Color contrast,
    Consequences on the lumen.

IV. Validation Tests

The aim of the validation stage was firstly to verify the consistency of the reasoning model and the knowledge representation with endoscopist behaviour and secondly to test the retrieval method using the endoscopist talks.

A. Material and method

Five short videoendoscopy sequences were submitted to 8 endoscopists: 4 seniors (10 years of practice at least) and 4 juniors (less than 3 years of practice). The videos sequences were selected according to the diagnostic difficulty and the characteristics of findings. A brief explanation on the project preceded each session but no information was delivered on the reasoning model or the knowledge representation. The 5 cases were always proposed in the same order: at first 3 simple cases and then 2 complex cases.

The endoscopists had to answer the following questions *always in the same order*:
    1. What are your diagnostic hypotheses?
    2. Which information elements have led to these conclusions?
    3. Have you ever encountered a similar case?

Endoscopists were encouraged to "think aloud" and the sessions were taped. They could ask questions to the investigator about information that was not present in the video, like the medical context. After a spontaneous talk period, the investigator asked questions to the endoscopist

to complete the scene description. The time was recorded: duration of spontaneous talk and total session, delay between the session start and the first endoscopist notice, the first diagnostic or context evocation, the correct diagnostic evocation. The discourse was matched with the knowledge base content in order to identify concordant descriptors. Two investigators controlled transcriptions. Two talk types were identified: fast (type 1) and delayed (type 2) diagnosis evocation according to the number of information before the first diagnostic hypothesis.

The measure of the computer-assisted diagnosis system performance was realized with indices, extracted from endoscopist talks. The system responses were compared with the correct endoscopical finding classes.

Statistical analysis was performed with ©SPSS software, using Log Rank test for censored data, Kruskall-Wallis ANOVA for quantitative data and exact Fisher test for qualitative data.

B.Results

The results concern 40 sessions for 8 endoscopists and 5 endoscopic cases. All descriptors types could be identified in the endoscopist spontaneous and total talk (tables 2, 3). Endoscopists mentioned 36 finding classes, 17 diagnosis classes and 16 distinct medical contexts, 28, 11 and 12 of them correctly matched with knowledge base content respectively. All 19 syntactic descriptors of endoscopical findings were mentioned, 17 of them spontaneously. The most frequent indices concerned anatomic location, shape, color and relief. Only 4 mentioned indices were not listed in the knowledge base. Type 1 talks were observed in 25 cases and type 2 in 15 cases. Similar last video situations had never been encountered in 3 cases: 2 juniors and 1senior.

TABLE 2.
Chronologic analysis of endoscopist spontaneous discourse.
Cross table of previous and next indices according to information type.

	Next evocation					
Previous evocation	A	B	C	D	E	F
A. Syntactic descriptors	163	30	1	19	30	
B. Endoscopic findings	29	17	2	3	15	
C. Medical context	1	1	7		5	1
D. Spatial relationships	14	6			2	
E. Endoscopic diagnosis	32	11	5	1	23	1
F. Diagnostic procedures				1		1
Session start	16	7	1		8	1
Total	**271**	**78**	**17**	**23**	**94**	**3**

The chronological analysis of spontaneous endoscopist talk (table 2) showed that most frequent information pairs consisted in: firstly the same information type listing; secondly combination between syntactic descriptors and endoscopic findings, diagnoses or spatial relationships; thirdly combination between endoscopic findings and diagnoses.

Comparison between simple and complex cases showed that the medical context and the diagnostic procedures were more often evocated in complex cases (table 3). The correct diagnosis was evoked in 21 of the 24 simple case sessions and 15 of 16 complex case sessions. All durations increased significantly in complex cases, except delay to correct diagnosis (table 4). Type-2 discourse was more frequent in complex cases (10/16 talks) than in simple cases (5/24 talks; p<0.05).

TABLE 3:
Means number of distinct information acoording to descriptor types and case comlexity. n: number of sessions; m: mean; sd: standard error; p:Kruskall-Wallis test; ns: not significant.

	Simple cases n=24		Complex cases n=16		
	m	sd	m	sd	p
Video play number	2.2	1.3	5.4	3.7	<0.01
Spontaneous talk					
Syntactic descriptors	5.5	3.7	5.1	2.8	ns
Endoscopic findings	1.6	1.2	1.5	1.4	ns
Medical context	0.1	0.3	0.8	1.0	<0.05
Spatial relationships	0.4	0.7	0.7	0.8	ns
Endoscopic diagnosis	1.5	0.9	1.8	1.1	ns
Diagnostic procedures	0.04	0.2	0.1	0.3	ns
Total talk					
Syntactic descriptors	8.3	3.0	9.6	3.4	ns
Endoscopic findings	3.2	1.8	3.6	1.7	ns
Medical context	1.7	1.4	2.7	1.7	<0.10
Spatial relationships	0.8	0.8	1.6	1.4	<0.10
Endoscopic diagnosis	2.2	1.1	3.1	2.0	ns
Diagnostic procedures	0.9	0.3	1.5	1.1	<0.05

.

TABLE 4.
Median time from session start to specific event according to case complexity. n: number of events; med: median time in seconds; sd: standard error; p: Log Rank test; ns: not significant.

	Simple cases 24 sessions			Complex cases 16 sessions			
	n	med	(sd)	n	med	(sd)	p
1^{rst} information	24	14	(4)	16	21	(2)	<0.05
1^{rst} diagnosis hypothesis	24	21	(2)	16	77	(18)	<0.0001
Correct diagnosis	21	29	(29)	15	112	(39)	ns
1^{rst} investigator request	24	78	(6)	16	145	(53)	<0.01
1^{rst} medical context	21	177	(26)	15	217	(62)	<0.05
Total duration	24	235	(28)	16	373	(66)	<0.001

Comparison according to experience showed that delay increased significantly for juniors. But, mobilized information was similar in juniors and seniors except the mean number of endoscopic diagnoses, which increased significantly in juniors ($3 \pm 1,8$ versus $2,1 \pm 1,1$; p<0,05). The correct diagnosis was evoked in 16/20 sessions for juniors and all sessions for seniors. The talk type frequency did not differ significantly between juniors and seniors (type 1: 12 senior and 13 junior sessions).

The 40 requests, built with indices extracted from endoscopist talks were submitted to the diagnostic aid system. They led to the correct finding class, as unique or first response in 25 cases, as second response in 5 cases and other rank response in 7. The correct finding was not found in 3 requests; all of them deriving from junior talks.

V. DISCUSSION

The computer-assisted diagnostic system has been built according to literature review and endoscopist expertise. The validation stage imposed to confront the reasoning model, the knowledge structure and content, and the research method to the reality of physician practice. This study was based on real endoscopic situations submitted to endoscopists with variable experience level. The "thinking aloud" method allows discerning the mean information mobilized during the diagnostic reasoning. While deliberately, no information was given on the conclusions of the previous stages of the project, so the talk was not influenced, at least during the spontaneous discourse. Because of the medical language richness (synonymy, paraphrase…), the discourse analysis implied an interpretation stage and could be influenced by the investigator thinking. To lower this potential bias risk, two investigators controlled and agreed on the talk transcription.

The distinction between endoscopic findings and endoscopic diagnosis is perceived in the endoscopist discourse, although the border can be ambiguous when the endoscopic finding constitutes the unique element of the endoscopic diagnosis. This ambiguity can explain the discourse chronology and the jumps from syntactic descriptors to endoscopic diagnosis. The chronological analysis suggests that the diagnostic approach is not only running a forward way from syntactic descriptors to endoscopic diagnosis, but admits many forward and backward motions between data and diagnostic hypothesis. This result can be partially explained by the study design and the discourse formulation but is in accordance with the reasoning processes that have been conceived at the previous stage: data-driven and hypothesis-driven approaches.

Fast diagnosis evocation (type 1) was most frequent in simple than in complex cases. This result suggests that when diagnostic problem is unusual, the endoscopist uses a more analytical approach based on syntactic descriptors, while faced to a common diagnostic problem, the perception is sudden according to an induction approach [6]. This result emphasizes the role of experience in the visual diagnostic process and the importance of syntactic descriptors in complex endoscopic problem solving.

No difference in talk type was found between juniors and seniors. This result can be explained by recent juniors experience. Although the last video session showed a very rare diagnostic case (duodenal metastasis), 2 of 4 juniors had encountered a similar case within one month before the test. The 2 other juniors never evoked the correct diagnosis. Only 1 senior had never encountered a similar case but evoked cautiously the correct diagnosis and proposed biopsies as a complementary diagnostic procedure.

The study confirms that the information mobilized during the endoscopic diagnostic process is consistent with the structure as well as the content of the knowledge base. The endoscopists introduce spatial relationships, most of them concerning part of the findings that is in accordance with the "object / sub-object" approach.

All types of objects and all syntactic findings descriptors have been mentioned by endoscopists. The absence of some diagnostic and finding classes was expected and does not modify the aid-system model: it just underlines the need for completing the knowledge base periodically.

The research method gives encouraging results while the correct response was found in 37 of 40 requests, and well ranked in 25. Further studies are needed to validate the request interface between the endoscopist and the retrieval process, in more realistic situation. Nevertheless, the present study suggests that it is nearly impossible for endoscopist to build an exhaustive request, especially when he is confronted with an unknown diagnostic situation. An interactive interface has yet been conceived for improving the user agreement: the most usual descriptors, as observed here, are the entry data and next, the research algorithm proposes other syntactic descriptors according to their aptitude to discriminate competing hypothesis.

Finally, the present study shows forward and backward reasoning in the endoscopic diagnostic process, while our computer-assisted diagnostic system runs only a data-driven approach. A further study is needed to integrate the 2 approaches in the diagnostic research system and to measure how these combined approaches could improve the responses.

REFERENCES

[1] J.M. Cauvin, C. Le Guillou, B. Solaiman, M. Robaszkiewicz, H. Gouérou, C. Roux. "Diagnostic reasoning in digestive tract endoscopy". *Proceedings of the BMES-EMBS congress,* Atlanta, 1999.

[2] J.M. Cauvin, C. Le Guillou, B. Solaiman, M. Robaszkiewicz , H. Gouérou. C. Roux. "Diagnostic reasoning by classification in upper digestive tract endoscopy". *Proceedings of the IEEE-EMBS congress*, Chicago, 2000.

[3] Z. Maratka. "Nomenclature illustrée en endoscopie digestive". Normed Verlag et Laboratoires Sandoz Eds. Homburg: 1992.

[4] R. Ford. "Competing conceptions of diagnostic reasoning – is there a way out ?". *Theor. Med. Bioeth,* vol. 19 pp. 59-72, 1998.

[5] M Crespi, M Delvaux, M Schapiro, C Venables, F Zwiebel. "Working party report of the Comittee for Minimal Standards of Terminology and Documentation in Digestive Endoscopy of the European Society of Gastrointestinal Endoscopy". *Am. J. Gastroenterol*, vol. 91, pp. 191-216, 1996.

[6] J Ridderikhoff. "Medical problem-solving : an exploration of strategies". *Med. Education,* vol. 25, pp. 196-207, 1991.